

LA-UR-03-3193

Approved for public release;  
distribution is unlimited.

c.1

Title: PROTEIN CRYSTALLOGRAPHY WITH SPALLATION  
NEUTRONS

Author(s): Paul A. Langan

Submitted to: Proceedings of International Symposium on Diffraction  
Structure Biology.  
Japan 2003  
Journal: Synchrotron Radiation



Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Form 836 (8/00)

59

# Protein Crystallography with Spallation Neutrons

Benno P. Schoenborn and Paul Langan

M888 Bioscience Division, Los Alamos National Laboratory, Los Alamos NM 87545, USA.  
E-mail: [langan\\_paul@lanl.gov](mailto:langan_paul@lanl.gov) Fax: +(505) 665 3204

Spallation neutron sources are ideal for diffraction studies of proteins and oriented molecular complexes. With spallation neutrons and their time dependent wavelength structure, one can select data with an optimal wavelength bandwidth and cover the whole Laue spectrum as time (wavelength) resolved diffraction data. This optimizes data quality with best peak to background ratios and provides spatial and energy resolution to eliminate peak overlaps. Such a Protein Crystallography Station (PCS) has been built and tested at Los Alamos Neutron Science Center. A partially coupled moderator is used to increase flux and data are collected by a cylindrical  $\text{He}^3$  detector covering  $120^\circ$  with 200mm height. The PCS is described along with examples of data collected from a number of proteins.

**Keywords:** Neutron diffraction; Protein Crystallography; enzyme mechanism; Protein Crystallography Station; PCS.

## 1. Introduction

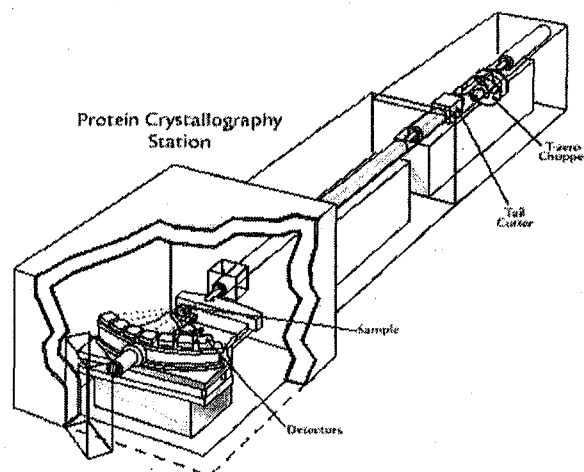
The PCS was built by Los Alamos National Laboratory's Bioscience Division at the Manuel Lujan Neutron Scattering Center, run by Los Alamos Neutron Science Center (LANSCE). Funded by the U.S. Department of Energy Office of Biological and Environmental Research, this station is the first PCS in the world to be built at a spallation neutron source, the first to use Time-of-Flight (TOF) methods, and the only neutron protein crystallography resource in North America. The PCS has been run as a user facility for the structural Biology community since August 2002. It provides an important proof of principle for future stations at advanced sources being planned and built worldwide.

Neutron diffraction is a powerful technique for locating hydrogen atoms and water molecules. Hydrogen atom positions and the coordination of water molecules cannot be directly determined using X-ray diffraction at resolutions typical for protein crystals. The majority of user experiments carried out on the PCS have been aimed at locating functionally important hydrogen atoms in order to determine detailed enzyme mechanism. However experiments aimed at investigating hydrogen bonding and hydration have also been carried out. In this report we give an overview of the PCS and then describe some of the preliminary data collected during the first year of operation as a user facility.

## 2. The Protein Crystallography Station

The PCS is located on Flight Path (FP) 15. Neutrons are produced by protons striking a W target in microsecond pulses with a frequency of 20Hz. These

high-energy neutrons are moderated in a reflector and then reduced to thermal energies in a moderator specifically tailored for protein crystallography (Schoenborn *et al*, 1999). The moderated neutrons are then extracted down a beam-pipe. A schematic representation of the beam layout on FP15 is shown in Figure 1. From the moderator neutrons travel a total flight path length of 28m down vacuum pipes with collimation inserts that taper the neutrons to produce a fine, almost parallel beam with a divergence of  $0.12^\circ$  that matches the mosaic of a typical protein crystal.



**Figure 1**  
A schematic layout of the PCS.

At 9.5m from the moderator the vacuum pipe is interrupted by two beam-shaping devices: a composite T0/T1 chopper and a proposed tail-cutting device (Langan *et al*, 2001). These devices remove unwanted high- and low-energy neutrons in each pulse, thus optimizing the neutron beam for high counting rates and low backgrounds at reasonable instrument resolutions. The vacuum pipe is tightly surrounded by heavy shielding until it reaches the sample position, where the shielding opens up to a large cave.

In the cave, a kappa-circle goniometer made by Huber moves the crystal through a number of orientations, recording a Laue pattern at each setting. A complete data set typically requires between 6-30 crystal settings depending on the symmetry of the crystal. An optional 8T super-conducting magnet can be mounted on the goniometer to orient membrane and fiber samples. A 700 series nitrogen cryostream cooler made by Oxford Cryosystems can be used to control the temperature of the sample down to  $\sim 100\text{K}$ . A large cylindrical PSD built by the Instrumentation Division of Brookhaven National Laboratory (BNL) collects as many of the spots as possible at each crystal setting without having to reposition the detector (Mahler *et al*, 1998).

The positions of the detected neutrons are calculated by advanced electronics associated with the detector and designed by BNL, before being passed to a VXI-based data-acquisition system designed by LANSCE where they are stored in memory along with their TOF. After a sufficient number of neutrons have been collected at a particular crystal and detector setting the data and the parameters that describe the instrument setting are written to a data archive in the form of a NeXus file (Klosowski *et al*, 1998).

The whole data collection process is coordinated from a Personal Computer (PC) with a Windows system (Microsoft) that runs a Java-based graphical user interface (GUI). A real time display of the data being collected on the detector is achieved on another PC that runs a LabVIEW (National Instruments Corporation) GUI to the detector electronics. On a third PC, a customized version of the commercial software package called d\*TREK (Molecular Structure Corporation) is used to display and process the archived NeXus files and ultimately to produce a measured intensity and an index for each diffraction spot. The intensities and indices provide the input required by standard protein-structure-refinement software and graphics programs that ultimately produce an image of the protein crystal's atomic structure.

### 3. Preliminary Results

#### 3.1 Human Insulin

The structure of Human insulin was recently determined by neutron diffraction using the BIX-3 protein crystallography station at the Japan Atomic Energy Research Institute (JAERI) (Maeda *et al*, 2003). Human insulin can crystallize in cubic space group, I2,3, which has 12 symmetry equivalent reflections. It is therefore an ideal system with which to test the corrections applied to symmetry equivalent reflections recorded at different wavelengths on the PCS. An experimental team from JAERI and Bioscience Division collected data for 16 hours at a single stationary setting as summarized in Table 1.

**Table 1**  
Data Collection parameters for Human Insulin.

Space Groups	I2,3
Unit cell dimensions (Å) <sup>a</sup>	a = 78.9
Wavelength range (Å)	0.6 - 5
Station	PCS
Temperature	295
Crystal settings	1
Average time per setting	16 hours
Observed reflections	4222
Observed reflections I/σ > 2	1490
Unique reflections I/σ > 2, d > 2.1 Å	1006

<sup>a</sup> Unit cell parameters are from X-ray diffraction.

The data were wavelength normalized and merged using d\*TREK. The value of  $R_{\text{merge}}$ , when calculated as a function of  $I/\sigma$ , rises rapidly for values of  $I/\sigma < 2$ , suggesting that below this threshold the intensities are not accurately measured, Table 2. When  $R_{\text{merge}}$  is tabulated as a function of resolution, after applying a  $I/\sigma > 2$  threshold, Table 3, significant data can be found out to ~1.8 Å resolution, although only 10% of the reflections lie beyond 2.1 Å. These results indicate that in order to collect a data set to 2.1 Å resolution, at least 10 crystal settings equivalent to 6 days of beam time would be required. The fact that a significant number of reflections are found out to resolutions of around 1.8 Å indicates that by extending the collection time at each crystal setting, an higher resolution data set could be obtained over a longer period of time.

**Table 2**  
 $R_{\text{merge}}$  versus  $I/\sigma$  for human insulin

$I/\sigma$	Num. Refs.	$\langle I/\sigma \rangle$	$R_{\text{merge}}$
18-20	393	59.4	0.008
16-18	49	17.0	0.146
14-16	29	14.9	0.179
12-14	58	13.2	0.243
10-12	77	11.0	0.130
8-10	102	8.9	0.148
6-8	154	6.8	0.182
4-6	219	4.8	0.295
2-4	749	2.7	0.503
0-2	2392	0.7	1.046

**Table 3**  
 $R_{\text{merge}}$ , completeness (%C), and  $\langle I/\sigma \rangle$  versus resolution for Human insulin. The cumulative  $R_{\text{merge}}$  and completeness are also given (Cum $R_{\text{merge}}$ , Cum%C). Only reflections with  $I/\sigma > 2$  are included in these calculations.

Res	No. Refs.	%C	Cum %C	$R_{\text{merge}}$	Cum $R_{\text{merge}}$	$\langle I/\sigma \rangle$
- 3.23	847	43.6	43.6	0.036	0.036	23.0
- 2.57	327	19.1	31.5	0.222	0.037	5.2
- 2.24	122	8.2	23.9	0.178	0.037	3.5
- 2.04	66	4.9	19.3	0.219	0.037	2.6
- 1.89	47	3.4	16.2	0.194	0.037	2.7
- 1.78	21	1.5	13.7	0.208	0.037	2.5

#### 3.2 Rubredoxin Mutant

Rubredoxin is a small protein that contains a single redox-active Fe(SCys)<sub>4</sub> center. Rubredoxin from *Pyrococcus furiosus*, a thermophilic archaeobacterium found near geothermal vents on the ocean floor, maintains its native structure at high temperatures (100°C). In order to study this unusual thermostability the hydrogen bonding pattern between the native protein and a series of mutants are being investigated using neutron diffraction. The neutron structures of the native protein and a triple mutant have already been determined using data collected on BIX-3 at JAERI (Chatake *et al*, 2002, Kurihara *et al*

2001). The experiment on the PCS concerns determining the neutron structure of a single mutant (Trp3 → Tyr3).

An experimental team from the University of Southern California, JAERI and Bioscience Division collected data at 9 crystal settings for 12 hours at each setting, over a period of 6 days, as summarized in Table 4. Data collection was restricted to 9 crystal settings because of lack of time. The data were wavelength normalized using LAUENORM (Campbell *et al.*, 1995). Only reflections with  $I/\sigma > 2$  were used for wavelength normalization and the resulting  $R_{\text{merge}}$  was 4.6%. The data were then merged using SCALA (Evans). Table 5 shows the values of  $R_{\text{merge}}$  and completeness as a function of resolution. The  $\sigma$  in Table 5 corresponds to the standard deviation in merging the reflections, not the error associated with measuring  $I$ . The fact that the value of  $I/\sigma$  tends to flatten off beyond  $\sim 2.5\text{\AA}$  suggests that this is about the limit of diffraction. The completeness at this limit is 65%, although the cumulative completeness is 74.8%.

**Table 4**  
Data Collection parameters for Rubredoxin single Mutant (Trp3 → Tyr3).

Space Groups	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell dimensions (Å) <sup>a</sup>	a = 78.9
Wavelength range (Å)	0.6 - 5
Station	PCS
Temperature	295
Crystal settings	9
Average time per setting	12 hours
Observed reflections	10315
Observed reflections $I/\sigma > 2$ , $d < 2.3\text{\AA}$	4065
Unique reflections $I/\sigma > 2$ , $d < 2.3\text{\AA}$	1252

<sup>a</sup> Unit cell parameters are from X-ray diffraction.

**Table 5**  
 $R_{\text{merge}}$ , completeness (%C), and  $\langle I/\sigma \rangle$  verses resolution for rubredoxin single mutant (Trp3 → Tyr3). The cumulative  $R_{\text{merge}}$  and completeness are also given (Cum $R_{\text{merge}}$ , Cum%C). Only reflections with  $I/\sigma > 2$  are included in these calculations.

Res	No. Refs.	%C	Cum %C	$R_{\text{merge}}$	Cum $R_{\text{merge}}$	$\langle I/\sigma \rangle$
- 7.27	173	87.3	87.3	0.025	0.025	24.7
7.27 - 5.14	373	86.5	86.8	0.147	0.033	4.0
5.14 - 4.20	426	83.2	85.2	0.162	0.035	4.0
4.20 - 3.64	464	80.3	83.4	0.188	0.037	3.1
3.64 - 3.25	476	74.6	80.9	0.235	0.039	2.4
3.25 - 2.97	465	71.5	78.6	0.307	0.040	1.5
2.97 - 2.75	435	70.4	77.0	0.344	0.040	1.7
2.75 - 2.57	404	65.0	74.8	0.506	0.041	1.1
2.57 - 2.42	427	65.5	73.3	0.510	0.041	1.1
2.42 - 2.30	422	63.0	71.8	0.515	0.042	1.1

### 3.1 Xylose Isomerase

The protein D-xylose isomerase is an enzyme that catalyzes the conversion of D-xylose to D-xylulose and glucose to fructose by hydrogen atom transfer. This

protein plays an important economic role in the food industry and is also the subject of intense research in other areas. One possible mechanism for the enzyme is metal ion-mediated ionization of water, the metal being magnesium. The catalytic motif of a magnesium cation, water molecule and enzyme carboxylate group, is common to many magnesium containing enzymes that are important in cancer research. By determining the exact catalytic mechanism important insights will be gained into a whole class of enzymes, with a view to designing better anticancer drugs.

An experimental team from Oak Ridge National Laboratory, Fox Chase Cancer Center, and Bioscience Division collected data at 23 crystal settings for on times ranging from 6 hours to 48 hours at each setting, as time became available (Hanson *et al.*, 2003). The average collection time at each setting was  $\sim 20$  hours, as shown in Table 6. Wavelength normalization with LAUENORM using only reflections with  $I/\sigma > 2$ , resulted in a value of 3.2% for  $R_{\text{merge}}$ . From Table 7, which shows the values of  $R_{\text{merge}}$ , completeness and  $\langle I/\sigma \rangle$  at a threshold of  $I/\sigma > 2$  as a function of resolution, the data set has a resolution of  $\sim 2.1\text{\AA}$ - $2.0\text{\AA}$ . Significant intensities were observed out to  $1.5\text{\AA}$  resolution and suggest that a higher resolution data set could be obtained by recording for longer.

**Table 6**  
Data Collection parameters for Xylose Isomerase.

Space Group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell dimensions (Å) <sup>a</sup>	a = 94, b = 100, c = 104
Wavelength range (Å)	0.6 - 5
Station	PCS
Temperature	295
Crystal settings	23
Average time per setting	20 hours
Observed reflections	181,797
Observed reflections $I/\sigma > 2$	62,692
Unique reflections $I/\sigma > 2$	24,931

<sup>a</sup> Unit cell parameters are from X-ray diffraction.

**Table 7**  
 $R_{\text{merge}}$ , completeness (%Comp), and  $\langle I/\sigma \rangle$  verses resolution for xylose isomerase. The cumulative  $R_{\text{merge}}$  and completeness are also given (Cum $R_{\text{merge}}$ , Cum%Comp). Only reflections with  $I/\sigma > 2$  are included in the above terms. The completeness for all reflections without a  $I/\sigma$  threshold, %AllComp is given for comparison.

Resolution	No. Refs.	%C	Cum %C	%AllC	$R_{\text{merge}}$	Cum $R_{\text{merge}}$
- 5.69	7,758	98.5	98.5	99.4	0.028	0.028
5.69 - 4.02	12,023	96.1	96.9	97.1	0.266	0.030
4.02 - 3.29	11,457	91.1	94.3	93.0	0.314	0.031
3.29 - 2.85	8,443	82.1	90.0	89.2	0.388	0.032
2.85 - 2.55	56,81	67.4	83.6	83.7	0.399	0.032
2.55 - 2.32	4,368	54.5	76.6	77.7	0.365	0.032
2.32 - 2.15	4,205	49.9	71.1	70.4	0.340	0.032
2.15 - 2.01	3,689	43.4	66.1	61.1	0.325	0.032
2.01 - 1.90	2,811	33.3	60.8	54.3	0.294	0.032
1.90 - 1.80	2,257	26.0	55.7	48.8	0.282	0.032

The PCS is funded by the Office of Science and the Office of Biological and Environmental Research of the U.S. Department of Energy.

## References

- Campbell, J.W., Habash, J., Helliwell, J.R., Moffat, K., *Information Quarterly for Protein Crystallography*, 18, July 1986, Daresbury Laboratory, UK.
- Chatake, T., Kurihara, K., Tanaka, I., Adams, M.W.W., Jenney, F.E.Jr., Tsyba, I. Bau, R. & Niimura, N. (2002), *Appl. Phys. A*, **74**[suppl], S1280-S1282.
- Evans, P., "SCALE", MRC Laboratory of Molecular Biology, Cambridge, UK.
- Hanson, B.L., Langan, P., Katz, A., Li, X., Harp, J., Glusker, J., Bunick, G. & Schoenborn, B.P. (2003) "A Preliminary Time-of-Flight Neutron Diffraction Study of *S. rubiginosus* D-Xylose Isomerase [EC5.3.1.5]" *Acta Cryst. D* (submitted).
- Klosowski, P. Koennecke, Tischler, J.Z. & Osborn, R. (1998), *Physica B*, **241-243**, 151.
- Kurihara, K., Tanaka, I., Adams, M.W.W., Jenny, F.E.Jr., Moiseeva, N. Bau, R. & Niimura, N. (2001) *J. Phys. Soc. Jpn. Suppl. Sect. A*, **70**, 400-402.
- Langan, P., Schoenborn, B.P. & Daemon, L.L. (2001) In *Neutron Optics*, James L. Wood & Ian Anderson, Eds, *Proceedings of SPIE* **4509**, 66-72.
- Maeda, M, Chatake, T., Tanaka, I., Osterman, A. & Niimura, N. (2003) "Crystallization of large single crystals of cubic insulin for neutron protein crystallography studies", submitted to *J. Synch.Rad.*, to be published as proceedings of ISDB 2003, Tsukuba, Japan.
- Mahler, G.J. Radeka, V., Schaknowski, N.A., Smith, G.C., Yu, B. & Zojceski, A. (1998), *IEEE Nuclear Science Symposium*.
- Schoenborn, B.P., Court, J.D., Larson, A.C. & Ferguson, P.J. (1999), *Neutrons Research* **7**, 89-106.